EM 1438

Re submit 3/12/04

# Document And Report Documentation Page Submitted as edoc_1075483867

| Report Documentation Page | Form Approved OMB No. 0704-0188 |
|---|---|

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

| 1. REPORT DATE **22 JUL 2003** | 2. REPORT TYPE **N/A** | 3. DATES COVERED **-** | |
|---|---|---|---|
| 4. TITLE AND SUBTITLE **Algebraic Equivalence of Conjugate Direction and Multistage Wiener Filters** | | 5a. CONTRACT NUMBER | |
| | | 5b. GRANT NUMBER | |
| | | 5c. PROGRAM ELEMENT NUMBER | |
| 6. AUTHOR(S) | | 5d. PROJECT NUMBER | |
| | | 5e. TASK NUMBER | |
| | | 5f. WORK UNIT NUMBER | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Office of Naval Research Supported by, CSU, MRC, SAIC** | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | 10. SPONSOR/MONITOR'S ACRONYM(S) | |
| | | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) | |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited** | | | |
| 13. SUPPLEMENTARY NOTES **Also see: ADM001520, ASAP Workshop 2003 , The original document contains color images.** | | | |
| 14. ABSTRACT | | | |
| 15. SUBJECT TERMS | | | |

| 16. SECURITY CLASSIFICATION OF: | | 17. | 18. | 19a. NAME OF RESPONSIBLE |
|---|---|---|---|---|

| a. REPORT unclassified | b. ABSTRACT unclassified | c. THIS PAGE unclassified | LIMITATION OF ABSTRACT UU | NUMBER OF PAGES 13 | PERSON Patricia Mawby, EM 1438 PHONE:(703) 767-9038 EMAIL:pmawby@dtic.mil |
|---|---|---|---|---|---|

pwd: cannot determine current directory!

# Algebraic Equivalence of Conjugate Direction and Multistage Wiener Filters

**Louis Scharf , Todd McWhorter , Edwin Chong ,
Scott Goldstein , Michael Zoltowski**

Supported by Office of Naval Research under various contracts to
CSU , MRC , SAIC

Program:

Define a very general class of *iterative subspace* Wiener filters.
Within this class, identify two interesting subclasses: *conjugate
direction* Wiener filters and *multistage* Wiener filters. Study their
equivalences.

Colorado State University

1

# Key Findings

1. Any subspace expansion in an iterative subspace Wiener filter that uses the direct sum of the past subspace and the current gradient turns out an orthogonal basis for the Krylov subspace.

2. For every multistage Wiener filter, there is a family of equivalent conjugate direction Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.

3. For every *orthogonal* multistage Wiener filter, there is a family of equivalent conjugate *gradient* Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.

4. If an *orthogonal* multistage Wiener filter is initialized at the cross-covariance, then it and its equivalent conjugate *gradient* Wiener filter turn out a Krylov subspace. And vice-versa.

# A Few (Among Many) Motivating Problems

**1. General linear channel**

$x \in C$

$y \in C^n$

$\hat{x}$

$w^*$

**2. Repeated measurements**

$y_i = x + n_i \quad \& \quad y = hx + n; \quad h^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$ : averaging vector

**3. Multisensor Array Processing**

$y_i = e^{jp^i}x + n_i \quad \& \quad y = hx + n; \quad h^T = \begin{bmatrix} 1 & e^{jp} & \cdots & e^{jp(n-1)} \end{bmatrix}$ : steering vector

**4. CDMA**

$y_i = h_i x + n_i \quad \& \quad y = hx + n; \quad h^T = \begin{bmatrix} h_1 & \cdots & h_n \end{bmatrix}$ : chipping vector

**Generally, $x$ is an unobserved complex scalar and $y$ is a measured vector that carries information about $x$, linearly.**

## Second-Order Characterization of The
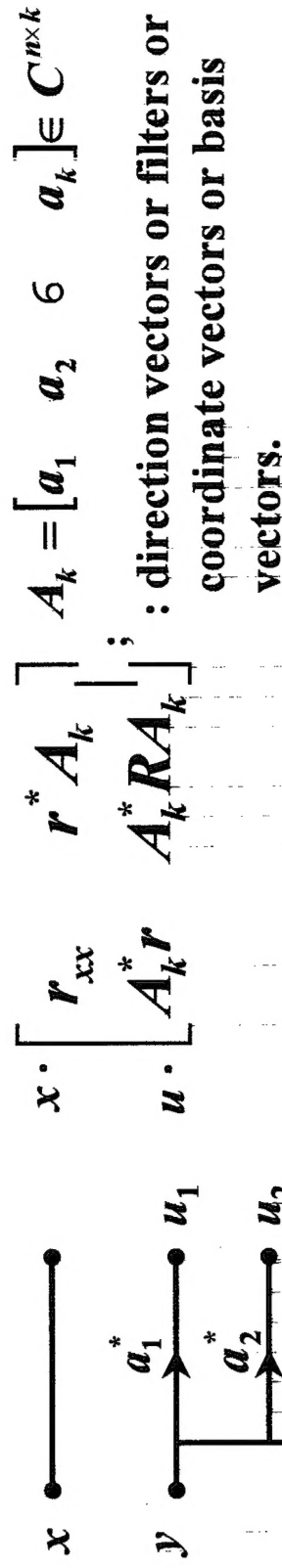## Two Channel Filtering Problem

$$x^\bullet \quad \begin{bmatrix} r_{xx} & r^* \\ r & R \end{bmatrix} \quad ; r_{xx} = Exx^* \in \mathbb{R}^+ \quad \& \quad r = Eyx^* \in C^n \quad \& \quad R = Eyy^* \in C^{n \times n}$$

$$y^\bullet$$



$$Q_{xx}(w) = r_{xx} - r^* w - w^* r + w^* R w$$

$$e \quad \begin{bmatrix} Q_{xx}(w) & -\gamma^*(w) \\ -\gamma(w) & R \end{bmatrix} \quad ; \quad : \text{MSE}$$

$$y$$

$$\gamma(w) = Rw - r : \quad gradient$$

For $\quad w = R^{-1} r, \quad \begin{bmatrix} r_{xx} - r^* R^{-1} r & 0 \\ 0 & R \end{bmatrix}.$

Thus, the plan is always to take the cov(x,y) to block diagonal form, where zeros reveal the orthogonality principle at work ($\gamma(w) = 0$).
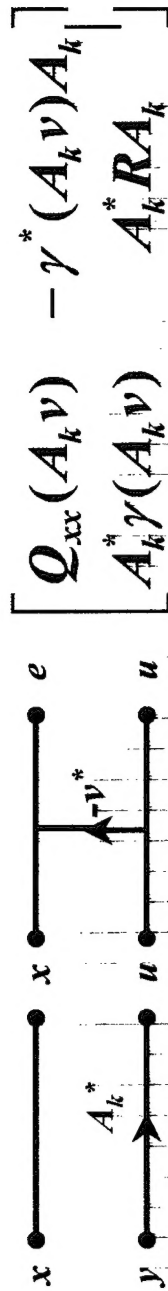
# The Expanding Subspace Idea: Resolve onto a Basis

$$x \cdot \begin{bmatrix} r_{xx} & r^* A_k \\ A_k^* r & A_k^* R A_k \end{bmatrix}$$

$$u \cdot \begin{bmatrix} \end{bmatrix}^{-1}$$

$A_k = [a_1 \quad a_2 \quad 6 \quad a_k] \in C^{n \times k}$

: direction vectors or filters or coordinate vectors or basis vectors.

$x$ •——————

$y$ •

$u_1$ •—— $a_1^*$ ↑

$u_2$ •—— $a_2^*$ ↑

7

$u_k$ •—— $a_k^*$ ↑

Think of this as a subspace expansion until the right tradeoff is achieved between bias and variance … in advance of detection, estimation, beamforming, spectrum analysis.

Generalize to matrix $a_i$ for vector $x$. We hope it generalizes to (true) filters $a_i(z)$ for time series.

# Filtering in the Expanding Subspace

**The original $(x, y)$ problem is now an $(x, u)$ problem:**

$$x \quad x \qquad\qquad e$$
$$A_k^* \quad {}_{\!-v^*} \qquad \begin{bmatrix} Q_{xx}(A_k v) & -\gamma^*(A_k v)A_k \\ A_k^* \gamma(A_k v) & A_k^* R A_k \end{bmatrix}$$
$$y \quad u \qquad\qquad u$$

**We are stuck with $A_k$, but we can optimize $v$. For the Wiener solution $v_k = (A_k^* R A_k)^{-1} A_k^* r$ $(w_k = A_k v_k)$, the covariance is**

$$\begin{bmatrix} r_{xx} - r^* A_k (A_k^* R A_k)^{-1} A_k^* r & 0 \\ 0 & A_k^* R A_k \end{bmatrix} \quad ; \quad 0 = A_k^* \gamma(A_k v_k)$$

**This bears comment.**

# Properties of the Expanding Subspace $\langle A_k \rangle$

In this construction call $A_k = [a_1 \quad a_2 \quad 5 \quad a_k]$ a basis for the k-dimensional subspace $\langle A_k \rangle$ and $u = A_k^* y$ the coordinates of $y$ in this basis. Call $\gamma_{k+1} = \gamma(A_k v_k)$ the gradient of $Q_{xx}(w)$ at the *subspace* solution $w_k = A_k v_k$, and $\Gamma_k = [\gamma_1 \quad \gamma_2 \quad 5 \quad \gamma_k]$ the matrix of gradients. Then $A_k^* \gamma_{k+1} = 0$ or $A_k^* \Gamma_k = C_k$ : lower triangular ... all a consequence of principle of orthogonality. The picture is this:

$\langle \gamma_{k+1} \rangle$

$\langle A_k \rangle$
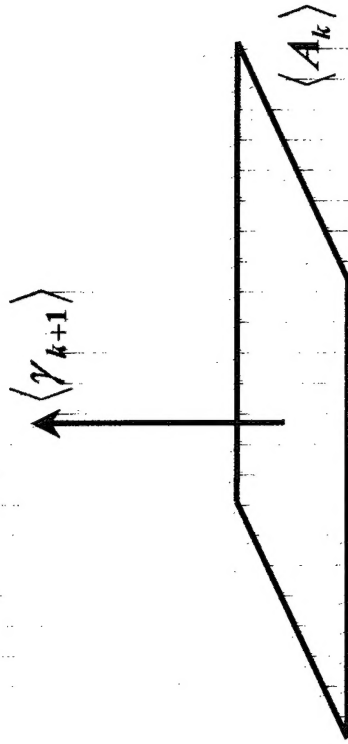
If $\langle A_k \rangle$ is initialized at $A_1 = a_1 = -\gamma_1 = r - Rw_0 = r$, then $[A_k, \gamma_{k+1}]$ turns out an orthogonal basis for the Krylov subspace $\langle K_{k+1} \rangle = \langle r, Rr, 5, R^k r \rangle$.

# Conjugate Directions and Multistage Wiener Filters

So far we have required nothing special of the filters or direction vectors $A_k$.

So let us now impose constraints

### CDWF

$A_k = D_k$ :  $re - name$

$D_k^* R D_k = \Sigma_k^2$ :  $diagonal$

$D_k^* \Gamma_k = C_k$ :  $lower\ triangular$

$(\langle D_k \rangle\ not\ generally\ Krylov)$

### MSWF

$A_k = G_k$ :  $re - name$

$G_k^* R G_k = T_k = B_k \Sigma_k^2 B_k^*$ :  $tri - diagonal$

$G_k^* \Gamma_k = B_k C_k$ :  $lower\ triangular$

$(\langle G_k \rangle\ not\ generally\ Krylov)$

CD $\longleftrightarrow$ MS

$D_k B_k^* = G_k$   or   $d_k b_{kk} + d_{k-1} b_{k-1k} = g_k$ :  $AR$

By constructing AR recursion between direction vectors, we go back and forth between CD and MS implementations.

# Conjugate Gradient & Orthogonal Multistage Wiener Filter

### CGWF

$A_k = D_k$ : $re-name$

$D_k^* R D_k = \Sigma_k^2$ : $diagonal$

$D_k^* \Gamma_k = C_k$ : $lower\ triangular$

$D_k B_k^* = \Gamma_k$ : $conjugate\ gradient$

$(\langle D_k \rangle\ generally\ Krylov)$

$\Gamma_k \Lambda_k = G_k$ : CG

### OMSWF

$A_k = G_k$ : $re-name$

$G_k^* R G_k = T_k = B_k \Sigma_k^2 B_k^*$ : $tri-diagonal$

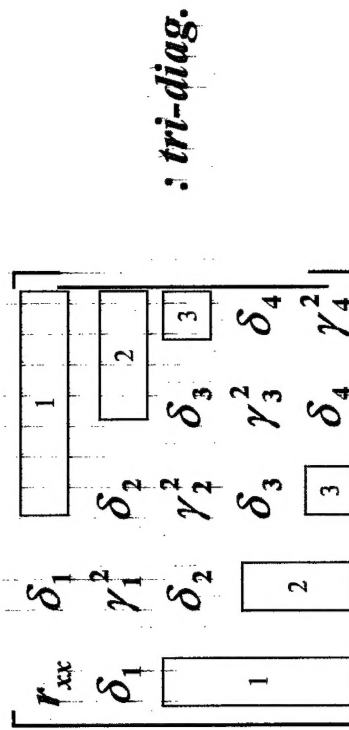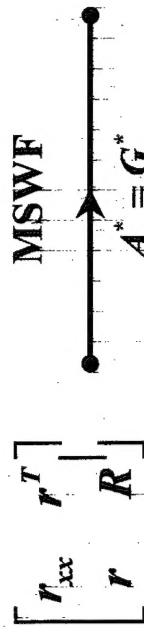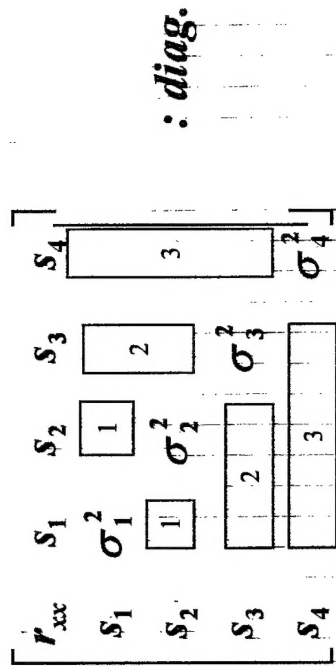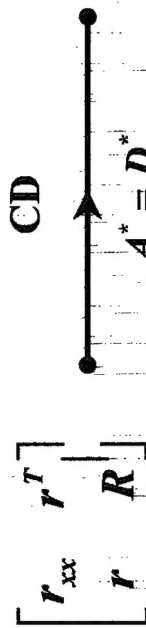$G_k^* \Gamma_k = B_k C_k$ : $lower\ triangular$

$G_k^* G_k = diagonal\ (orthogonal)$

$(\langle G_k \rangle\ generally\ Krylov)$

OMS : $G_k = \Gamma_k$

$\longleftrightarrow$

$D_k B_k^* = \Gamma_k$ or $G_k$ : $AR$

By constructing AR recursion between direction vectors, we go back and forth between CG and OMS implementations.

# A Little More Detail about Implementation of CD and MS Wiener Filters

In the analysis stage, the order in which zeros are forced into the covariance matrices is this

$$\begin{bmatrix} r_{xx} & r^T \\ r & R \end{bmatrix}$$

**CD**

$$A^* = D^*$$



: diag.

**MSWF**

$$\begin{bmatrix} r_{xx} & r^T \\ r & R \end{bmatrix}$$

$$A^* = G^*$$



: tri-diag.

The zero pattern of CD makes direction vectors *R-conjugate*. The zero pattern of MSWF makes direction vectors *R-tridiagonal*. In the case of CGWF and OMSWF, the *multistage vectors* of the OMSWF are identical to the *gradient vectors* of the CGWF.

In their synthesis steps CDWF and MSWF take their analyzed covariance matrices to diagonal form, where the orthogonality condition is enforced. Therefore, they have the same minimum MSE:

$$\begin{bmatrix} r_{xx} & s^T \\ s & \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & 6 & 0 \\ 0 & \sigma_2^2 & 0 & 7 \\ 7 & 0 & \sigma_3^2 & 0 \\ 0 & 6 & 0 & \sigma_4^2 \end{bmatrix} \xrightarrow{\ \ CD\ \ } \begin{bmatrix} Q_{xx} & 0^T : \text{ortho} \\ 0 & \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & 6 & 0 \\ 0 & \sigma_2^2 & 0 & 7 \\ 7 & 0 & \sigma_3^2 & 0 \\ 0 & 6 & 0 & \sigma_4^2 \end{bmatrix}$$

$$\begin{bmatrix} r_{xx} & \delta_1 & 0 & 6 & 0 \\ \delta_1 & \gamma_1^2 & \delta_2 & 0 & 7 \\ 0 & \delta_2 & \gamma_2^2 & \delta_3 & 0 \\ 7 & 0 & \delta_3 & \gamma_3^2 & \delta_4 \\ 0 & 6 & 0 & \delta_4 & \gamma_4^2 \end{bmatrix} \xrightarrow{\ \ MSWF\ \ } \begin{bmatrix} Q_{xx} & 0^T : \text{ortho} \\ 0 & \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_1^2 & 0 & 6 & 0 \\ 0 & \hat{\gamma}_2^2 & 0 & 7 \\ 7 & 0 & \hat{\gamma}_3^2 & 0 \\ 0 & 6 & 0 & \hat{\gamma}_4^2 \end{bmatrix}$$

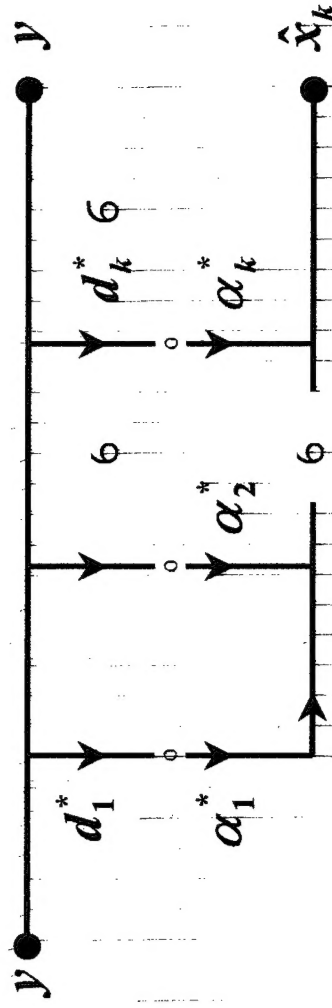This proves equivalence of CDWF and MSWF. Their respective MMSE formulas are

$$Q_{xx} = r_{xx} - \frac{s_1^2}{\sigma_1^2} - 6 - 7 - \frac{s_k^2}{\sigma_k^2} = r_{xx} - \cfrac{\delta_1^2}{\gamma_1^2 - \cfrac{\delta_2^2}{\gamma_2^2 - \cfrac{\delta_3^2}{5 - \gamma_k^2}}}$$

: continued sum vs. continued fraction

The picture is this.

**CDWF:**

$y \quad d_1^* \quad \circ \quad 6 \quad \circ \quad d_k^* \quad 6 \quad y$

$\alpha_1^* \quad \alpha_2^* \quad \alpha_k^*$

$\hat{x}_k$

**MSWF:**

$y \quad g_1^* \quad 6 \quad g_k^* \quad 6 \quad y$

$\beta_{1k}^* \quad \beta_{2k}^* \quad \beta_{kk}^*$

$\hat{x}_k$
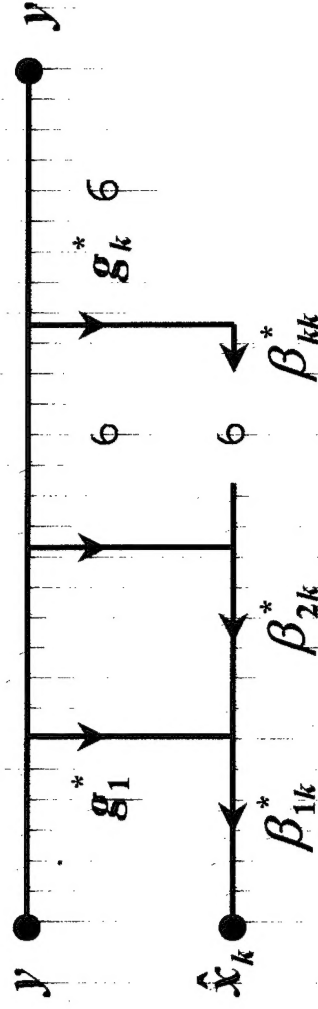
For an optimization theorist, a *line search* for the optimum *step size* $\alpha_i$ or $\beta_i$ in direction $d_i$ or $g_i$ is equivalent to the filtering theorist's MMSE *weight computation* in the synthesis stage of the iterative filter.

## Conclusions

1. Any subspace expansion in an iterative subspace Wiener filter that uses the direct sum of the past subspace and the current gradient turns out an orthogonal basis for the Krylov subspace.

2. For every multistage Wiener filter, there is a family of equivalent conjugate direction Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.

3. For every *orthogonal* multistage Wiener filter, there is a family of equivalent conjugate *gradient* Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.

4. If an *orthogonal* multistage Wiener filter is initialized at the cross-covariance, then it and its equivalent conjugate *gradient* Wiener filter turn out a Krylov subspace. And vice-versa.

5. All of this generalizes to matrix filters $a_i$. Does it generalize to time series filters $a_i(z)$? To something like least squares filtering on manifolds, as opposed to (linear) subspaces? To kernel-based nonlinear processing of measurements?